

XX WP1: 1997-394558/37.
XX
XX Preparation of S-layer proteins by expressing sbs-A gene in Gram
XX negative bacterium - or new sbs-B gene in any host, also new
XX recombinant proteins containing heterologous inserts, e.g.
XX epitope(s), useful as vaccines and adjuvants
XX
XX Claim 1: Pages 9-14; 31pp; German.
XX
XX The present sequence encodes the Bacillus stearothermophilus PV72
XX S-layer protein, sbs-A. S-layer structures can be used as vaccines
XX or adjuvants, particularly when they include a bacterial ghost that
XX may contain additional epitopes in its membrane. Other uses of
XX recombinant sbs-A, depending on the nature of the inserted peptide,
XX are as an universal carrier for biotinylated reactants for use in
XX immunological or hybridisation assays (the insert is streptavidin),
XX to induce immune responses (epitopes), as a reagent for removing
XX cytokine or toxin from serum (antigenic epitopes), as a molecular
XX spinning nozzle (polyhydroxybutyrate synthase) and as a molecular
XX laser (luciferase).
XX
XX Sequence 3687 BP: 1316 A; 660 C; 715 G; 996 T; 0 other;

Query Match 100.0%; Score 3687; DB 18; Length 3687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggtatgagaaagagctgtgnaactgaacagcaagcgctatttcacaaagtcattt 60
Db 1 atggtatgagaaagagctgtgnaactgaacagcaagcgctatttcacaaagtcattt 60
QY 61 gtgcgtgcacaaatccaaacgctctgaaagcgctacagatgtgacaaacagtagtaaccaa 120
Db 61 gtgcgtgcacaaatccaaacgctctgaaagcgctacagatgtgacaaacagtagtaaccaa 120
QY 121 gcaaaagcacagttcaaaaagcatctactactacagccatacagtaacggaacatgtt 180
Db 121 gcaaaagcacagttcaaaaagcatctactactacagccatacagtaacggaacatgtt 180
QY 181 gattccccaacttaacagatgtatagtctgaatatacaaaaacggaaaaacgatccct 240
Db 181 gattccccaacttaacagatgtatagtctgaatatacaaaaacggaaaaacgatccct 240
QY 241 gatcggtacagctttagtaataaagcaggtggcggaataaagacgcttacttagctgt 300
Db 241 gatcggtacagctttagtaataaagcaggtggcggaataaagacgcttacttagctgt 300
QY 301 ttacaaaagaatatagaacttaagtttcaaaagcaaaccttaaatctggcgaaagctgt 360
Db 301 ttacaaaagaatatagaacttaagtttcaaaagcaaaccttaaatctggcgaaagctgt 360
QY 361 gtaagcaactataatgaatgcttacaactatgcacaacaaattgaagaaatggcccaag 420
Db 361 gtaagcaactataatgaatgcttacaactatgcacaacaaattgaagaaatggcccaag 420
QY 421 ctgaagcgctgcttccaacaaagatttagaaaagcgagaaacatfactacaaatatt 480
Db 421 ctgaagcgctgcttccaacaaagatttagaaaagcgagaaacatfactacaaatatt 480
QY 481 cctttgaatataaaactgcacagtcatttagatcggtatattgylaaaacaaactcgt 540
Db 481 cctttgaatataaaactgcacagtcatttagatcggtatattgylaaaacaaactcgt 540
QY 541 gattacttcgcttacaatttaaaacaaagacagaactcgcgcagcgttaatttat 600
Db 541 gattacttcgcttacaatttaaaacaaagacagaactcgcgcagcgttaatttat 600
QY 601 gatattacggttcacatgaagcgcgcggaagtaacaagacgctgtgaaagcagcaattta 660
Db 601 gatattacggttcacatgaagcgcgcggaagtaacaagacgctgtgaaagcagcaattta 660

QY 661 gacaaagctaaagctgctgttatacaaatcaatcaactactacccaagaatgacagatgt 720
Db 661 gacaaagctaaagctgctgttatacaaatcaatcaactactacccaagaatgacagatgt 720
QY 721 ttcaaaacgtaactaaacgaataagtagcgaataaagcatttgatgacagatgaactgtgctt 780
Db 721 ttcaaaacgtaactaaacgaataagtagcgaataaagcatttgatgacagatgaactgtgctt 780
QY 781 actccaagaatgaaagttaagtcgatatcaactcaacaaacaaagctgtgaattaaac 840
Db 781 actccaagaatgaaagttaagtcgatatcaactcaacaaacaaagctgtgaattaaac 840
QY 841 gcaatracaggtacaggaacactaaatatacaacttcagctgtgcgaataaagataca 900
Db 841 gcaatracaggtacaggaacactaaatatacaacttcagctgtgcgaataaagataca 900
QY 901 gtaaacgtaataactgtacgtatctataaagggcggaacattccattcccttaatt 960
Db 901 gtaaacgtaataactgtacgtatctataaagggcggaacattccattcccttaatt 960
QY 961 acggtacagatgttcttatactacagacggaacaaactacatgtgagcttcaactcca 1020
Db 961 acggtacagatgttcttatactacagacggaacaaactacatgtgagcttcaactcca 1020
QY 1021 ttccgaataataacaggaatataaagtagtagttaaaggtatataagacacaaatgacaa 1080
Db 1021 ttccgaataataacaggaatataaagtagtagttaaaggtatataaagacacaaatgacaa 1080
QY 1081 gattttaaagaagatgcatcttaacgcttgcgaataagagctgttagttactcaag 1140
Db 1081 gattttaaagaagatgcatcttaacgcttgcgaataagagctgttagttactcaag 1140
QY 1141 ttctgaactaatgttaacaaacaaacactctctgaacttagcagcaggttccctcagact 1200
Db 1141 ttctgaactaatgttaacaaacaaacactctctgaacttagcagcaggttccctcagact 1200
QY 1201 gacgaacttaacagcagatattgataagctgttgaacacctgaacactgttaacagctcg 1260
Db 1201 gacgaacttaacagcagatattgataagctgttgaacacctgaacactgttaacagctcg 1260
QY 1261 aacgttactatacagatgttgaactgaacaaacgcatcccgtaattgcactactctt 1320
Db 1261 aacgttactatacagatgttgaactgaacaaacgcatcccgtaattgcactactctt 1320
QY 1321 ggttcacaaactacattacgttcaaaagagcgttagtaactgtgtaaacataataaact 1380
Db 1321 ggttcacaaactacattacgttcaaaagagcgttagtaactgtgtaaacataataaact 1380
QY 1381 gctatcaaatatgttaaaacattaaactgtgtacaatgacagaagcttaacagatgtgtc 1440
Db 1381 gctatcaaatatgttaaaacattaaactgtgtacaatgacagaagcttaacagatgtgtc 1440
QY 1441 actgcaaaagcctacagcaacactgttgcacgctccctcaactttgggtgtacaaat 1500
Db 1441 actgcaaaagcctacagcaacactgttgcacgctccctcaactttgggtgtacaaat 1500
QY 1501 ttatctacgtgttcttcaacaaacaaatgtttgggtaaatgtgcgggtgtgtgatgaa 1560
Db 1501 ttatctacgtgttcttcaacaaacaaatgtttgggtaaatgtgcgggtgtgtgatgaa 1560
QY 1561 gctggaacttatattcctgtgtctcaatatacaacaaacgcttgcataagttagaacaa 1620
Db 1561 gctggaacttatattcctgtgtctcaatatacaacaaacgcttgcataagttagaacaa 1620
QY 1621 tctacttagctgataactttgatttagttgaaagaatctggtacaggtttgtctct 1680
Db 1621 tctacttagctgataactttgatttagttgaaagaatctggtacaggtttgtctct 1680
QY 1681 gaactaaataataatgcagagcctaaatgtgaactttaggtgcgaagcgaagccttaa 1740
Db 1681 gaactaaataataatgcagagcctaaatgtgaactttaggtgcgaagcgaagccttaa 1740
QY 1741 gaaataacatctatcaaatcaaaattaaaagagcttgaagttccgataaagtaattga 1800

Dh	1741	gaaataacatactataccaatccaataatcaaaaaagccttgaagcgcgataaaggtattgaa	1800
Qy	1801	tttaggcacggtttaaagagaaaaacataatgaatttcaaaaactaaagactaaactgtcctaca	1860
Dh	1801	tttaggcacggtttaaagagaaaaacataatgaatttcaaaaactaaagactaaactgtcctaca	1860
Qy	1861	gtctattagcgttaacgltctcaaaaaatggcgcgcgcgtgatataaagtaactgaagctcaagaa	1920
Dh	1861	gtctattagcgttaacgltctcaaaaaatggcgcgcgcgtgatataaagtaactgaagctcaagaa	1920
Qy	1921	tttaactgtgaagtctccacagagaaatttaaatatccttaaatgtctcaaacggtttgggtgac	1980
Dh	1921	tttaactgtgaagtctccacagagaaatttaaatatccttaaatgtctcaaacggtttgggtgac	1980
Qy	1981	acaatcaacaatagcgtcnaagtttgcgtgaatgaataaagcgggttcaaacctatctgtcctaca	2040
Dh	1981	acaatcaacaatagcgtcnaagtttgcgtgaatgaataaagcgggttcaaacctatctgtcctaca	2040
Qy	2041	gcaagtgacacataatctccagctagttgttgaagcgggttactgtgcataagatgaaacatacaa	2100
Dh	2041	gcaagtgacacataatctccagctagttgttgaagcgggttactgtgcataagatgaaacatacaa	2100
Qy	2101	gtgaaaggtgcgtgcgttaaccaaattagaacgtaaaccaagggttacaacattagatggttcggt	2160
Dh	2101	gtgaaaggtgcgtgcgttaaccaaattagaacgtaaaccaagggttacaacattagatggttcggt	2160
Qy	2161	aaaaggtgcacacgcgtcgtttaaagaatgctgcacaaatgcaaatcctttagcaactaatcat	2220
Dh	2161	aaaaggtgcacacgcgtcgtttaaagaatgctgcacaaatgcaaatcctttagcaactaatcat	2220
Qy	2221	atctatatacttacaacactgaagagttcaagacgttaacagcaacaaacggttacaanaagtaac	2280
Dh	2221	atctatatacttacaacactgaagagttcaagacgttcaagacgttacaacaaacggttacaanaagtaac	2280
Qy	2281	aaaaggtgatcttctttaaagaacgctgtgatgcagcttaactacacttcgaaacggttgatgcaggt	2340
Dh	2281	aaaaggtgatcttctttaaagaacgctgtgatgcagcttaactacacttcgaaacggttgatgcaggt	2340
Qy	2341	caaaaaattcaactacccaatttttagcgaagaaattaaaaactctatgttgttttttagtgtgt	2400
Dh	2341	caaaaaattcaactacccaatttttagcgaagaaattaaaaactctctatgttgttttttagtgtgt	2400
Qy	2401	ggcnaaagttaactgttcgagaaatttaacaaacaaacgcgaatggttaatggtttactcgtgaca	2460
Dh	2401	ggcnaaagttaactgttcgagaaatttaacaaacaaacgcgaatggttaatggtttactcgtgaca	2460
Qy	2461	actgtatcagctgcgtccccaagacagatgcaaatgtgtaaaagttaacagctgctgtgtttaca	2520
Dh	2461	actgtatcagctgcgtccccaagacagatgcaaatgtgtaaaagttaacagctgctgtgtttaca	2520
Qy	2521	ttcaacgcggtcttbaaataaagcnaaagatgagaaatttggctctgtgtagtaagtaagttct	2580
Dh	2521	ttcaacgcggtcttbaaataaagcnaaagatgagaaatttggctctgtgtagtaagtaagttct	2580
Qy	2581	tctactgataagaaattgcgtgaatgagatgtgttaagtgaatlaagcaaaaaagatatattat	2640
Dh	2581	tctactgataagaaattgcgtgaatgagatgtgttaagtgaatlaagcaaaaaagatatattat	2640
Qy	2641	cgttacaacagcgtgagacacactgttagctctgtgnaagcgtcgtgtgaacaaagatggt	2700
Dh	2641	cgttacaacagcgtgagacacactgttagctctgtgnaagcgtcgtgtgaacaaagatggt	2700
Qy	2701	caaaaacgctcttgcgtgcatcttcccaaaaagaagcgtcaattgatatacaataagaagcttatta	2760
Dh	2701	caaaaacgctcttgcgtgcatcttcccaaaaagaagcgtgcatatgatacnaactaaagaagcttatta	2760
Qy	2761	gtctgaattcaatgaaaaactgaatttagcggaaagttaaacctgagacaactcgtgtttaagaat	2820
Dh	2761	gtctgaattcaatgaaaaactgaatttagcggaaagttaaacctgagacaactcgtgtttaagaat	2820
Qy	2821	gagcaaggttaatgvcggttagcgtgactgttaacagcatlaagcgtttcaaaaataaattt	2880

Db	2821	gcagcaggttaatgctggtagctggtgtagctggtgaacgcgaatgaagaggtttctcaataaatatt	2880
Qy	2881	gtatctactccatctccagaattataaagctggtgacagtttactctgttaacaattgacggt	2940
Db	2881	gtatctactccatctccagaattataaagctggtgacagtttactctgttaacaattgacggt	2940
Qy	2941	gtgagaggttaagtaggttaacacaaatctcctaaatacattacttcgtttcaagcttatct	3000
Db	2941	gtgagaggttaagtaggttaacacaaatctcctaaatacattacttcgtttcaagaggttatct	3000
Qy	3001	gcgagatccaaaggttatcttcaatccagcatggtgcgagcgtgcgaagtttaagcttgacggtct	3060
Db	3001	gcgagatccaaaggttatcttcaatccagcatggtgcgagcgtgcgaagtttaagcgttgacggtct	3060
Qy	3061	aaacaattatcaatltgaaattccagcgattcagttcccaacccaacaatctcatcttlaagaag	3120
Db	3061	aaacaattatcaatltgaaattccagcgattcagttcccaacccaacaatctcatcttlaagaag	3120
Qy	3121	gctgaacggaactctcatcttaactaattacactttgataaagtataatgataaataaaca	3180
Db	3121	gctgaacggaactctcatcttaactaattacactttgataaagtataaataaataaaca	3180
Qy	3181	tacaaaattgattctcccaagaaggtgataacactggaagagtttaacataataagttagaac	3240
Db	3181	tacaaaattgattctcccaagaaggtgataacactggaagagtttaacataataagttagaac	3240
Qy	3241	gttccaagaagttcttcaaaactggtgactgtaattgtatagccaagtttaacttccatccacggt	3300
Db	3241	gttccaagaagttcttcaaaactggtgactgtaattgtatagccaagtttaacttccatccacggt	3300
Qy	3301	tctgtgtgctactgtgacgaagaataaacctgctctatgtagagcggtgtgtcatggaatggaaca	3360
Db	3301	tctgtgtgctactgtgacgaagaataaacctgctctatgtagagcggtgtgtcatggaatggaaca	3360
Qy	3361	agctatactcagaggtgtgtgtggaagcaaacacgactcggctcgttagctgacttcgtttgcgag	3420
Db	3361	agctatactcagaggtgtgtgtggaagcaaacacgactcggctcgttagctgacttcgtttgcgag	3420
Qy	3421	ccaggtgccttccatctctccagaaggtatcgatttaacgaattgcaactgtgacagtaaca	3480
Db	3421	ccaggtgccttccatctctccagaaggtatcgatttaacgaattgcaactgtgacagtaaca	3480
Qy	3481	aataattactgtgtgaataaactgltgaaagtattttccaagaagaggtgtagacgcagacacat	3540
Db	3481	aataattactgtgtgaataaactgltgaaagtattttccaagaagaggtgtagacgcagacacat	3540
Qy	3541	gatgcagaggtgctactaagaagagacattgataataacacagttactctctttagtactggtat	3600
Db	3541	gatgcagaggtgctactaagaagagacattgataataacacagttactctctttagtactggtat	3600
Qy	3601	aacgcacagagactttaagattggttgaaggtggaggtttaagaatgcacagcaagtaattgttga	3660
Db	3601	aacgcacagagactttaagattggttgaaggtggaggtttaagaatgcacagcaagtaattgttga	3660
Qy	3661	gatactataatcatcttatataagtaa	3687
Db	3661	gatactataatcatcttatataagtaa	3687
RESULT	2	:	
X22748			
ID	X22748	standard; DNA; 3687 bp.	
XX			
AC	X22748:	?	
XX			
DT	24-AUG-1999	(first entry)	
XX			
DE	B	steatothermophilus sbas DNA.	
KM	sbas protein; S-layer protein; Gram-negative; prokaryotic host cell;		
KM	integration; cytoplasmic membrane; secretion; periplasmic space; toxin;		
KM	eukaryotic host cell; vaccine; adjuvant; immunogenic epitope; luciferase		
KM	immunostimulant; cytokine; poly(vi-droxybutyrate); pib synthase; body fluid;		

MM molecular laser; universal carrier molecule; monomolecular layer; ds.
XX
OS Bacillus stearothermophilus.
XX
FH key Location/Qualifiers
FT 1..3687
FT /tag- a
FT /product- "sbsa"
FT sig-peptide 1..90
FT /tag- b
FT mat-peptide 91..3684
FT /tag- c
XX
PN DE19732829-A1.
XX
PD 04-FEB-1999.
XX
PR 30-JUL-1997; 97DE-1032829.
XX
PR 30-JUL-1997; 97DE-1032829.
XX
PA (LDBL/) LDBITZ W.
XX
PI Lubitz W, Resch S;
XX
DR WPI: 1999-122189/11.
XX P-PSDB: W93252.
XX
PT Producing S-layer proteins in Gram-negative bacteria or eukaryotes -
PT integrated into membranes or organelles or secreted into periplasma
PT or growth medium, and nucleic acid encoding S-layer proteins with
PT peptide insertions, used in vaccines or for enzymatic reactions
XX
PS Claim 5, Page 9-14; 34pp; German.
XX
CC This invention describes a method for the production of a S-layer protein
CC (I) which comprises (a) preparing a Gram-negative prokaryotic host cell
CC transformed with nucleic acid (II) encoding (I), linked to a signal
CC sequence (SS) that encodes a protein which causes at least one of (i)
CC integration of (I) into the external or cytoplasmic membranes and/or
CC (ii) secretion of (I) into the periplasmic space or extracellular medium,
CC (b) culturing the cell to express (I) and (c) optionally recovering (I)
CC from the membranes, periplasmic space and/or extracellular medium.
CC Alternatively, a eukaryotic cell is used as host and then the SS, which
CC is optional, promotes integration into the cytoplasmic membrane or an
CC organelle and/or secretion into the extracellular medium. (I) and
CC derived structures, may include a wide variety of polypeptide inserts and
CC are useful as (i) vaccines or adjuvants (with immunogenic epitopes or
CC immunostimulants) inserts such as cytokines (ii) as reactants (inserts
CC are enzymes, e.g. polyhydroxybutyrate (PHB) synthase for use as a
CC 'molecular spinnerette' for production of PHB or luciferase for use as
CC molecular laser (when combined with substrate and oxygen) and (iii) as
CC universal carrier molecule (streptavidin is inserted) for use in
CC hybridisation and immunoassays, or for selective elimination of
CC cytokines, toxins etc. from body fluids (inserts are specific binding
CC epitopes). In this system, heterologous (I) do not form inclusion bodies
CC but rather monomolecular layers, and in eukaryotic cells they undergo
CC glycosylation. This sequence encodes the Bacillus stearothermophilus
CC sbsa protein which is used to illustrate the method of the invention.
XX
SQ Sequence 3687 BP; 1316 A; 650 C; 715 G; 996 T; 0 other;

Query Match 100.0%; Score 3687; DB 20; Length 3687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggataggaaaaagctgtgaacacagacagcaagtcctatcgagaaagtcattt 60
Db 1 atggataggaaaaagctgtgaacacagacagcaagtcctatcgagaaagtcattt 60
QY 61 gtgcgtgcaaatccaaagcgtctgaagcggtacagagtgtagcaacagtagtaagccaa 120
Db 1141 gtgcgtgcaaatccaaagcgtctgaagcggtacagagtgtagcaacagtagtaagccaa 120

Db 61 gtgcgtgcaaatccaaagcgtctgaagcggtacagagtgtagcaacagtagtaagccaa 120
QY 121 gcaaaagcagcttcaaaaagacatactacttaagcattcaataagaaagtaagtcgt 180
Db 121 gcaaaagcagcttcaaaaagacatactacttaagcattcaataagaaagtaagtcgt 180
QY 181 gaattcccaacaattaaagatataatgctgaatacaacaagaagaaagacataacct 240
Db 181 gaattcccaacaattaaagatataatgctgaatacaacaagaagaaagacataacct 240
QY 241 gatgcggtagcaatagtgaaataagcagtgccgcaaaaagacgcttacttagctgat 300
Db 241 gatgcggtagcaatagtgaaataagcagtgccgcaaaaagacgcttacttagctgat 300
QY 301 ttcaaaaagaatataatgaatctgtttcaaaagcaaaccttaattctgacgaagtcgt 360
Db 301 ttcaaaaagaatataatgaatctgtttcaaaagcaaaccttaattctgacgaagtcgt 360
QY 361 gtacgaacttcaactgcttgaacacatgcacacaaattagacgaaatgcgccaagag 420
Db 361 gtacgaacttcaactgcttgaacacatgcacacaaattagacgaaatgcgccaagag 420
QY 421 ctagaagcgctgttcaagaacaaagatttagaanaagcagaacaatactacacaaatt 480
Db 421 ctagaagcgctgttcaagaacaaagatttagaanaagcagaacaatactacacaaatt 480
QY 481 ccttatgaatttaaaactgcacagcattttgacgcggtatattgtaaaaactcgt 540
Db 481 ccttatgaatttaaaactgcacagcattttgacgcggtatattgtaaaaactcgt 540
QY 541 gatttactcgcctcattttaaagaagaacagaacgaacttcgcagacgttaattat 600
Db 541 gatttactcgcctcattttaaagaagaacagaacgaacttcgcagacgttaattat 600
QY 601 gatattacggtgcaatgaaagcgcggaagatgacaaagcgtgtgaagcaggaattta 660
Db 601 gatattacggtgcaatgaaagcgcggaagatgacaaagcgtgtgaagcaggaattta 660
QY 661 gaaacagctaaagctcgtttgatacaatcaatcaatcttaccaaaagtaaaagtcgt 720
Db 661 gaaacagctaaagctcgtttgatacaatcaatcaatcttaccaaaagtaaaagtcgt 720
QY 721 ttcaaaactgtaacaaagaatgagcaaaagaacttagatgcagatgaagtcgctt 780
Db 721 ttcaaaactgtaacaaagaatgagcaaaagaacttagatgcagatgaagtcgctt 780
QY 781 actccaaagctgaaaggttgaagtcgatacaacacccaaacaaagcgtttgaattaca 840
Db 781 actccaaagctgaaaggttgaagtcgatacaacacccaaacaaagcgtttgaattaca 840
QY 841 gcaatccagtgaaagcagacaaatacaacttcaagcttcgcgcgaatgaaagtcata 900
Db 841 gcaatccagtgaaagcagacaaatacaacttcaagcttcgcgcgaatgaaagtcata 900
QY 901 gtaaacgtaaatctgtaagtatacttaaaagtgaaagtaaaccttccattcccttaatt 960
Db 901 gtaaacgtaaatctgtaagtatacttaaaagtgaaagtaaaccttccattcccttaatt 960
QY 961 acggcagatgcttcttactatcacaagcagcaaaaactatactcgtgtgaactccaa 1020
Db 961 acggcagatgcttcttactatcacaagcagcaaaaactatactcgtgtgaactccaa 1020
QY 1021 ttccgaaataatagcagagataaagtagttaaaggtataaagacaaaatgacaaa 1080
Db 1021 ttccgaaataatagcagagataaagtagttaaaggtataaagacaaaatgacaaa 1080
QY 1081 gaatttaaaagagatgcattcacttcaagcttcgaaagatgcgtttagttactcaagtg 1140
Db 1081 gaatttaaaagagatgcattcacttcaagcttcgaaagatgcgtttagttactcaagtg 1140
QY 1141 ttggaactaatgtlaacaacaacacttctgttaaaacttagcagcaggtacttgcgact 1200
Db 1141 ttggaactaatgtlaacaacaacacttctgttaaaacttagcagcaggtacttgcgact 1200

QY 1201 gacgatacttaacagtagatttgtaagttgttaagcacctgaacacgtgaacagctcg 1260
|||||
Db 1201 gaagatacttaacagtagatttgtaagttgttaagcacctgaacacgtgaacagctcg 1260
QY 1261 aacgttacttaccagaatgtgaacacgtgaacacgttcacgaatctgaatctacttct 1320
1261 aacgttacttaccagaatgtgaacacgttcacgaatctgaatctacttct 1320
Db 1261 aacgttacttaccagaatgtgaacacgttcacgaatctgaatctacttct 1320
QY 1321 ggttctacaaatcacttaagtttaaaagagcttagtaactggttaacaaataaactt 1380
|||||
Db 1321 ggttctacaaatcacttaagtttaaaagagcttagtaactggttaacaaataaactt 1380
QY 1381 gcatcaataatgtttaaataacatgaatcgttcaatgaagaagcttaagaatggttttc 1440
|||||
Db 1381 gcatcaataatgtttaaataacatgaatcgttcaatgaagaagcttaagaatggttttc 1440
QY 1441 actgcgaacacgtacacgaacacacgtgtgtacacgctcactactacttaagttgtataact 1500
1441 actgcgaacacgtacacgaacacacgtgtgtacacgctcactactacttaagttgtataact 1500
Db 1441 actgcgaacacgtacacgaacacacgtgtgtacacgctcactactacttaagttgtataact 1500
QY 1501 ttaactactgtgtctcttaacaaataatgttggtgaataatggtgtgtgtgtgtgaatgaa 1560
|||||
Db 1501 ttaactactgtgtctcttaacaaataatgttggtgaataatggtgtgtgtgtgtgaatgaa 1560
QY 1561 gctgcgaactattatctctgttcaatcacaacacgcttctgaactgaagttaagtgagaa 1620
|||||
Db 1561 gctgcgaactattatctctgttcaatcacaacacgcttctgaactgaagttaagtgagaa 1620
QY 1621 tctacttcaagtgaacttctgtattatgtgaataaagaatctgtgaactgtgtgtctct 1680
1621 tctacttcaagtgaacttctgtattatgtgaataaagaatctgtgaactgtgtgtctct 1680
Db 1621 tctacttcaagtgaacttctgtattatgtgaataaagaatctgtgaactgtgtgtctct 1680
QY 1681 gaacttaataataatgtgcagacgttaaaatgttaacttagtgcgaagaagcgagacttaaa 1740
1681 gaacttaataataatgtgcagacgttaaaatgttaacttagtgcgaagaagcgagacttaaa 1740
Db 1681 gaacttaataataatgtgcagacgttaaaatgttaacttagtgcgaagaagcgagacttaaa 1740
QY 1741 gaataataataatcacttaataacataatgaagaagcgttggaactgaagaatgaa 1800
|||||
Db 1741 gaataataataatcacttaataacataatgaagaagcgttggaactgaagaatgaa 1800
QY 1801 tttaggcactgttaacgagaaacataatgagttcnaaaactgaagaactgaactgtcctaca 1860
1801 tttaggcactgttaacgagaaacataatgagttcnaaaactgaagaactgaactgtcctaca 1860
Db 1801 tttaggcactgttaacgagaaacataatgagttcnaaaactgaagaactgaactgtcctaca 1860
QY 1861 gttattatgagtaacgcttaaaatgtgcagacgttgatataaagttaactgaactgaaga 1920
1861 gttattatgagtaacgcttaaaatgtgcagacgttgatataaagttaactgaactgaaga 1920
Db 1861 gttattatgagtaacgcttaaaatgtgcagacgttgatataaagttaactgaactgaaga 1920
QY 1921 tttaactgtgaagttctcgaagaatttaaatcatttaactgaacacgcttcgagctgagc 1980
1921 tttaactgtgaagttctcgaagaatttaaatcatttaactgaacacgcttcgagctgagc 1980
Db 1921 tttaactgtgaagttctcgaagaatttaaatcatttaactgaacacgcttcgagctgagc 1980
QY 1981 acaatacacatacgttcgaagtgtgtgtaaaagcggtgcgaactatctgtcctaca 2040
1981 acaatacacatacgttcgaagtgtgtgtaaaagcggtgcgaactatctgtcctaca 2040
Db 1981 acaatacacatacgttcgaagtgtgtgtaaaagcggtgcgaactatctgtcctaca 2040
QY 2041 gcaagtgaacatcttcgaagcttagttgtaagcggttaactggtcaagaatggaacatacaaa 2100
2041 gcaagtgaacatcttcgaagcttagttgtaagcggttaactggtcaagaatggaacatacaaa 2100
Db 2041 gcaagtgaacatcttcgaagcttagttgtaagcggttaactggtcaagaatggaacatacaaa 2100
QY 2101 gtttaaggttgctgtcaaaccaattgaagaatgaagaaggtgaagaatgaagttgttcgt 2160
2101 gtttaaggttgctgtcaaaccaattgaagaatgaagaaggtgaagaatgaagttgttcgt 2160
Db 2101 gtttaaggttgctgtcaaaccaattgaagaatgaagaaggtgaagaatgaagttgttcgt 2160
QY 2161 aaggtgcgaacagctcctgtttaagaatgtgtgcgaatgaatacttaggaacttaactat 2220
2161 aaggtgcgaacagctcctgtttaagaatgtgtgcgaatgaatacttaggaacttaactat 2220
Db 2161 aaggtgcgaacagctcctgtttaagaatgtgtgcgaatgaatacttaggaacttaactat 2220
QY 2221 attataacatttcaaacatgtgaaggtcaagaatgaagaatgaagaatgaagaatgaagaat 2280
2221 attataacatttcaaacatgtgaaggtcaagaatgaagaatgaagaatgaagaatgaagaat 2280
Db 2221 attataacatttcaaacatgtgaaggtcaagaatgaagaatgaagaatgaagaatgaagaat 2280

QY 2281 aaggtgatctcttaaaagacgtgtgaatgcagttactaacttaacgaacgtgtgaatgcaggt 2340
2281 aaggtgatctcttaaaagacgtgtgaatgcagttactaacttaacgaacgtgtgaatgcaggt 2340
Db 2281 aaggtgatctcttaaaagacgtgtgaatgcagttactaacttaacgaacgtgtgaatgcaggt 2340
QY 2341 caaaaattcactatccaaatttagcgaagaatttaaaaactctggtgttctttagtggt 2400
2341 caaaaattcactatccaaatttagcgaagaatttaaaaactctggtgttctttagtggt 2400
Db 2341 caaaaattcactatccaaatttagcgaagaatttaaaaactctggtgttctttagtggt 2400
QY 2401 ggcgaagttacgttcgagaatattaaacaaacaggttggttaagtctgttactgtgaaca 2460
2401 ggcgaagttacgttcgagaatattaaacaaacaggttggttaagtctgttactgtgaaca 2460
Db 2401 ggcgaagttacgttcgagaatattaaacaaacaggttggttaagtctgttactgtgaaca 2460
QY 2461 actgtatcagttgtctccttaagagaagatgcgaatgttaaaatgaacagctgtgtgtaca 2520
2461 actgtatcagttgtctccttaagagaagatgcgaatgttaaaatgaacagctgtgtgtaca 2520
Db 2461 actgtatcagttgtctccttaagagaagatgcgaatgttaaaatgaacagctgtgtgtaca 2520
QY 2521 tcaactgtctctgaacaaatgaacaaagatgtgaatgtgtgtgtgtgtgtgtgtgtgtgt 2580
2521 tcaactgtctctgaacaaatgaacaaagatgtgaatgtgtgtgtgtgtgtgtgtgtgtgt 2580
Db 2521 tcaactgtctctgaacaaatgaacaaagatgtgaatgtgtgtgtgtgtgtgtgtgtgtgt 2580
QY 2581 tctactgtatgtgaattgt 2640
2581 tctactgtatgtgaattgt 2640
Db 2581 tctactgtatgtgaattgt 2640
QY 2641 cgtttacaacacgctgcgagacacacactgtgactctcgtgtgaagctgtgtgcgaagaatggt 2700
2641 cgtttacaacacgctgcgagacacacactgtgactctcgtgtgaagctgtgtgcgaagaatggt 2700
Db 2641 cgtttacaacacgctgcgagacacacactgtgactctcgtgtgaagctgtgtgcgaagaatggt 2700
QY 2701 caaaacgctctgt 2760
2701 caaaacgctctgt 2760
Db 2701 caaaacgctctgt 2760
QY 2761 gttgaatttcaatgaagaatgt 2820
2761 gttgaatttcaatgaagaatgt 2820
Db 2761 gttgaatttcaatgaagaatgt 2820
QY 2821 gcgacggtatgcgt 2880
2821 gcgacggtatgcgt 2880
Db 2821 gcgacggtatgcgt 2880
QY 2881 gtttcaactcactcgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 2940
2881 gtttcaactcactcgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 2940
Db 2881 gtttcaactcactcgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 2940
QY 2941 gttgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatga 3000
2941 gttgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatga 3000
Db 2941 gttgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatga 3000
QY 3001 gcgaatcgaacgtttactcacttaacgaacatgtgtgcagctgaacgtgtgcagctgtc 3060
3001 gcgaatcgaacgtttactcacttaacgaacatgtgtgcagctgaacgtgtgcagctgtc 3060
Db 3001 gcgaatcgaacgtttactcacttaacgaacatgtgtgcagctgaacgtgtgcagctgtc 3060
QY 3061 aaaaacaaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatga 3120
3061 aaaaacaaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatga 3120
Db 3061 aaaaacaaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatga 3120
QY 3121 gctgcaggaactcacttacttaacttaacttaacttaacttaacttaacttaacttaactta 3180
3121 gctgcaggaactcacttacttaacttaacttaacttaacttaacttaacttaacttaactta 3180
Db 3121 gctgcaggaactcacttacttaacttaacttaacttaacttaacttaacttaacttaactta 3180
QY 3181 tacaataatgttactcacaagaatgtgaacacgttggaaggtttactcacaatgtgtgtgaca 3240
3181 tacaataatgttactcacaagaatgtgaacacgttggaaggtttactcacaatgtgtgtgaca 3240
Db 3181 tacaataatgttactcacaagaatgtgaacacgttggaaggtttactcacaatgtgtgtgaca 3240
QY 3241 gtttcaaaagatttcaaaactgtgtactgtatgtatgtatgtatgtatgtatgtatgtatgt 3300
3241 gtttcaaaagatttcaaaactgtgtactgtatgtatgtatgtatgtatgtatgtatgtatgt 3300
Db 3241 gtttcaaaagatttcaaaactgtgtactgtatgtatgtatgtatgtatgtatgtatgtatgt 3300
QY 3301 tctgtgtgtactgtgcgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 3360
3301 tctgtgtgtactgtgcgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 3360
Db 3301 tctgtgtgtactgtgcgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 3360
QY 3361 agctatactcagatgtgtgcgaacacgacactcgttctgtgtgtgtgtgtgtgtgtgtgtgt 3420
3361 agctatactcagatgtgtgcgaacacgacactcgttctgtgtgtgtgtgtgtgtgtgtgtgt 3420

Db	3361	agctatactcaggaatgtctgcagcaacacgaactcgcgtctgttcgctgacttcgttcgagag	3420
Qy	3421	ccagttgcocctcaactatctcagaagaatctacgaatttaacgaatgcacattgtgcagtaaca	3480
Db	3421	ccagttgcocctcaactatctcagaagaatctacgaatttaacgaatgcacattgtgcagtaaca	3480
Qy	3481	aattttcctgctgtgttaaaactgtttgaagttatctttcaaaaagaagaagtgtgagcgaacacat	3540
Db	3481	aattttcctgctgtgttaaaactgtttgaagttatctttcaaaaagaagaagtgtgagcgaacacat	3540
Qy	3541	gattccaggtgtctactaaaggagacattagtaatttaacacagttaccctcttagtaattgat	3600
Db	3541	gattccaggtgtctactaaaggagacattagtaatttaacacagttaccctcttagtaattgat	3600
Qy	3601	aacagcagaagacttaataagattgttgtaaagtgtgaatlaaagaatgcagcagtaattgtgca	3660
Db	3601	aacagcagaagacttaataagattgttgtaaagtgtgaatlaaagaatgcagcagtaattgtgca	3660
Qy	3661	gatactattatactctctattataagtaa	3687
Db	3661	gatactattatactctctattataagtaa	3687

RESULT	3
ID	T08695 standard; DNA; 3832 BP.
XX	T08695
XX	T08695,
XX	15-JUL-1996 (first entry)
XX	sbsA gene encoding S-layer protein.
XX	sbsA gene; S-layer; cell surface layer; expression; ss
XX	Bacillus stearothermophilus.
XX	
XX	Key
XX	Location/Qualifiers
XX	37..3723
XX	/*tag= a
XX	19..25
XX	RBS
XX	/*tag= b
XX	1..36
XX	/*tag= c
XX	3723..3832
XX	/*tag= d
XX	37..126
XX	/*tag= e
XX	/note= "clalm 1"
XX	127..3720
XX	/*tag= f
XX	/note= "clalm 3"
XX	3769..3806
XX	stem_loop
XX	/*tag= g
XX	/note= "terminator"
XX	
XX	DE4425527-A1.
XX	
XX	25-JAN-1996.
XX	
XX	19-JUL-1994; 94DE-4425527.
XX	
XX	19-JUL-1994; 94DE-4425527.
XX	
XX	(VOGE-) VOGELBUSCH GMBH.
XX	
XX	Lablitz W;
XX	
XX	WPI: 1996-077933/09.
XX	
XX	p-PSDE: 87673.
XX	
XX	Nucleic acid encoding signal peptide of Bacillus stearothermophilus

PT S-layer protein - which has a lysine content of at least 10 per
 pr cent.
 xx
 xx
 PS Disclosure; Page 6-7; 12pp; German.
 xx
 xx
 CC The claimed signal peptide encoding sequence (see tag e of T08695)
 CC is pref. operably linked at the 3' end to a protein encoding
 CC sequence. The protein is pref. the S-layer protein (see tag f of
 CC T08695). At the 5' end, the signal peptide encoding sequence is
 CC pref. linked to an expression control sequence, pref. the sequence
 CC given in T08695.
 CC
 CC
 SQ Sequence 3832 BP; 1355 A; 686 C; 749 G; 1042 T; 0 other;

Query Match	100.0%	Score 3687	DB 17	Length 3832
Best Local Similarity	100.0%	Pred. No. 0		
Matches 3687	Conservative	0	Mismatches	0
			Indels	Gaps
			0	0
QY 1 atggaatggaataaagcgtgtgaactagcaaacagcagtgctatgccaagtgcatc 60				
Db 37 atggatggaataaagcgtgtgaactagcaaacagcagtgctatgccaagtgcatc 96				
QY 61 gtccgtgtgcacatcacaacgcgtctctgaagcgcgtgtacagatgtgacaaagtaagccaa 120				
Db 97 gtccgtgtgcacatcacaacgcgtctctgaagcgcgtgtacagatgtgacaaagtaagccaa 156				
QY 121 gcaaaaagcgcgtttcaaaaaagcacaatactattacagcacaacgttaaagcgaactgtt 180				
Db 157 gcaaaaagcgcgtttcaaaaaagcacaatactattacagcacaacgttaaagcgaactgtt 216				
QY 181 gaattcccaaacatcaacgattgatagtgtatcgtgatacaaaaagcggaaaaaacgataccgt 240				
Db 217 gaattcccaaacatcaacgattgatagtgtatcgtgatacaaaaagcggaaaaaacgataccgt 276				
QY 241 gatgcgtgtgacataatgataaagaacaggtggcgcggaaaaaagacgcttacttaagctgat 300				
Db 277 gatgcgtgtgacataatgataaagaacaggtggcgcggaaaaaagacgcttacttaagctgat 336				
QY 301 ttaaaaaaagaataatgaataacttaagcttttccaagcacaaccccaaatctctgcaagctcgt 360				
Db 337 ttaaaaaaagaataatgaataacttaagcttttccaagcacaaccccaaatctctgcaagctcgt 396				
QY 361 gttagcaactataatcgtatgacttaccacatacgaaacaaaatttagacgaaatggccaaagat 420				
Db 397 gttagcaactataatcgtatgacttaccacatacgaaacaaaatttagacgaaatggccaaagat 456				
QY 421 cttagagcgtctgtctcaagcacaagaattttagaaaaaagcagaacaatcattatcacaataat 480				
Db 457 cttagagcgtctgtctcaagcacaagaattttagaaaaaagcagaacaatcattatcacaataat 516				
QY 481 ccttatgaaataaataaacccgcacagctcatcttaagtgcggtatattgtttaaaccaactcgt 540				
Db 517 ccttatgaaataaataaacccgcacagctcatcttaagtgcggtatattgtttaaaccaactcgt 576				
QY 541 gatttaactcgtctcacaattttaaagcacaagaagcacaagaacttcgcgcacagccttaatttt 600				
Db 577 gatttaactcgtctcacaattttaaagcacaagaagcacaagaacttcgcgcacagccttaatttt 636				
QY 601 gatatataccgtttgcataatgaaagcgcgcgaagtacaaagacgctgtgtgaagccagcacaattta 660				
Db 637 gatatataccgtttgcataatgaaagcgcgcgaagtacaaagacgctgtgtgaagccagcacaattta 696				
QY 661 gatacaagctataagctgcgtgtgtgatacaaatcaatcaatacttccacaagaatcaacagatcgt 720				
Db 697 gatacaagctataagctgcgtgtgtgatacaaatcaatcaatacttccacaagaatcaacagatcgt 756				
QY 721 ttcaaaactgtacataacaggaatgacgaaaaaagcattagatgtgcagatgtgaagctcgcgt 780				
Db 757 ttcaaaactgtacataacaggaatgacgaaaaaagcattagatgtgcagatgtgaagctcgcgt 816				
QY 781 actccaagaagtgtgaagtgtgaatgtcgatacaatcacaacacaagaagcgtgtgtgaatttaaca 840				

QY	3001	gcgcatccaacgcttatctcttcataacacgcttcgcgaagcgcgagcttaagcttgaacgcttcct	3050
Db	3037	gcgaatccaacgcttatctcttcataacacgcttcgcgaagcgcgagcttgaacgcttcct	3096
QY	3061	aaacaacatacaatgatcatcagcgatctcagttctccaacaccacaatacactcttaagag	3120
Db	3097	aaacaacatacaatgatcatcagcgatctcagttctccaacaccacaatacactcttaagag	3156
QY	3121	gctgaacgaaccttcatttactaacattccactttagtaaatgttaataatgaaataaaca	3180
Db	3157	gctgaacgaaccttcatttactaacattccactttagtaaatgttaataatgaaataaaca	3216
QY	3181	tacaaatttatattccccaagaagtgtaaaccttgacagagattactccaataatagacttaaca	3240
Db	3217	tacaaatttatattccccaagaagtgtaaaccttgacagagattactccaataatagacttaaca	3276
QY	3241	gttcccaaaagatttccaacctgtgactcgtatattgacagcaaaagttaacattccaccagt	3300
Db	3277	gttcccaaaagatttccaacctgtgactcgtatattgacagcaaaagttaacattccaccagt	3336
QY	3301	tctgtgtcactcgaacgaagataaaacctgctctagtagagcgctgtgtcattgnaatgnaaca	3360
Db	3337	tctgtgtcactcgaacgaagataaaacctgctctagtagagcgctgtgtcattgnaatgnaaca	3396
QY	3361	agcctatctcaagatctgcgcgaacaaacgaactcggctctgtatgcgcgacttcgttcgagag	3420
Db	3397	agcctatctcaagatctgcgcgaacaaacgaactcggctctgtatgcgcgacttcgttcgagag	3456
QY	3421	ccaaattcccttcacattctcagaaggttatccgatttagcgaatgcgaactgttcagtagtaaca	3480
Db	3457	ccaaattcccttcacattctcagaaggttatccgatttagcgaatgcgaactgttcagtagtaaca	3516
QY	3481	aataattcagatgataaaacctgtttgaagttattttccaagaagagtgtagacgcgaacacat	3540
Db	3517	aataattcagatgataaaacctgtttgaagttattttccaagaagagtgtagacgcgaacacat	3576
QY	3541	gattcagagtggtctctcctaagaagacattagtaataacaacagttaccctcttagtacttgat	3600
Db	3577	gattcagagtggtctctcctaagaagacattagtaataacaacagttaccctcttagtacttgat	3636
QY	3601	aaacgcagaagccttaataagattgtttgaagtgtgaaggttaagaatgcgcgaagtgtaattgca	3660
Db	3637	aaacgcagaagccttaataagattgtttgaagtgtgaaggttaagaatgcgcgaagtgtaattgca	3696
QY	3661	gatactattacattcatatataaagtaa 3687	
Db	3697	gatactattacattcatatataaagtaa 3723	
RESULT	4		
X78246	ID	X78246 standard; DNA; 3768 BP.	
XX	XX	X78246;	
AC	XX		
XX	XX		
D7	XX	24-AUG-1999 (first entry)	
DE	XX	B. stearothermophilus sbAa/Bacteriophage fd Gene 3 fusion protein DNA.	
KW	XX	sbAa protein; S-layer protein; Gram-negative; prokaryotic host cell;	
KW	XX	integration; cytoplasmic membrane; secretion; periplasmic space; toxin;	
KW	XX	eukaryotic host cell; vaccine; adjuvant; immunogenic epitope; luciferase;	
KW	XX	immunostimulant; cytokine; polyhydroxybutyrate; PHB synthase; body fluid;	
KW	XX	molecular laser; universal carrier molecule; monomolecular layer; Gene 3	
KW	XX	fusion gene; ds.	
XX	XX		
OS	XX	Synthetic.	
OS	XX	Bacteriophage fd.	
OS	XX	Bacillus stearothermophilus	
PN	XX	DE19732829-AL.	
XX	XX		
DD	XX	04-FEB-1999.	

XX 30-JUL-1997; 97DE-1032829.
PE 30-JUL-1997; 97DE-1032829.
PR 30-JUL-1997; 97DE-1032829.
XX
PA (LUBITZ/) LUBITZ W.
XX
PI Lubitz W, Resch S;
DR WP1; 1999-122189/11.
XX
P1 Producing S-layer proteins in Gram-negative bacteria or eukaryotes -
P1 integrated into membranes or organelles or secreted into periplasma
P1 or growth medium, and nucleic acid encoding S-layer proteins with
PT peptide insertions, used in vaccines or for enzymatic reactions
XX
PS Example 8; Fig 5; 34pp; German.

This invention describes a method for the production of a S-layer protein (I) which comprises (a) preparing a Gram-negative prokaryotic host cell transformed with nucleic acid (II) encoding (I), linked to a signal sequence (SS) that encodes a protein which causes at least one of (I) integration of (I) into the external or cytoplasmic membranes and/or (II) secretion of (I) into the periplasmic space or extracellular medium, (b) culturing the cell to express (I) and (c) optionally recovering (I) from the membranes, periplasmic space and/or extracellular medium. Alternatively, a eukaryotic cell is used as host and then the SS, which is optional, promotes integration into the cytoplasmic membrane or an organelle and/or secretion into the extracellular medium. (I), and derived structures, may include a wide variety of polypeptide inserts and are useful as (i) vaccines or adjuvants (with immunogenic epitopes or immunostimulants inserts such as cytokines) (ii) as reactors (inserts are enzymes, e.g. polyhydroxybutyrate (PHB) synthase for use as a 'molecular spinnereite' for production of PHB or luciferase for use as molecular laser (when combined with substrate and oxygen)) and (iii) as universal carrier molecule (streptavidin is inserted) for use in hybridisation and immuno assays, or for selective elimination of cytokines, toxins etc. from body fluids (inserts are specific binding epitopes). In this system, heterologous (I) do not form inclusion bodies but rather monomolecular layers, and in eukaryotic cells they undergo glycosylation. This sequence encodes a fusion gene constructed from the Bacillus stearotherophilus sbxa gene and the bacteriophage fd Gene 3 signal sequence which is used in the method of the invention.

Db 319 ggcgaagctcgtgtagaacttacatcgatcgcttacacaactatgcacaaatagacgaa 378
QY 409 atggcgcaaaagctagagcgctcgtctcaagcaaaagatttagaaaaagcagaaataac 468
Db 379 atggcgcaaaagctagagcgctcgtctcaagcaaaagatttagaaaaagcagaaataac 438
QY 469 taccacaagaattcccttagaataaacaacgcagacagatctatcttagatccgtataggt 528
Db 439 taccacaagaattcccttagaataaacaacgcagacagatctatcttagatccgtataggt 498
QY 529 aaaaacaactcgttagtacttactcgtctacatcttaagaacaaagacacagaaattcgcgac 588
Db 499 aaaaacaactcgttagtacttactcgtctacatcttaagaacaaagacacagaaattcgcgac 558
QY 589 agattaattatgatattacacgtttgcgaatgaagcgcgcgaagacacagacgcgtgtgaa 648
Db 559 agattaattatgatattacacgtttgcgaatgaagcgcgcgaagacacagacgcgtgtgaa 618
QY 649 gcggcgcaatttagacaagaagctgctgtgtgatcaaatcaatcaacttaccacaaa 708
Db 619 gcggcgcaatttagacaagaagctgctgtgtgatcaaatcaatcaacttaccacaaa 678
QY 709 gtaacagatgcttctcaaaaactgaaactgaagtagcgaaanaaagcattagatgcagat 768
Db 679 gtaacagatgcttctcaaaaactgaaactgaagtagcgaaanaaagcattagatgcagat 738
QY 769 gaagcgtcgcttactccaaaagtctgaaagtgtaagtgcgttaacatctcaaaaagaagct 828
Db 739 gaagcgtcgcttactccaaaagtctgaaagtgtaagtgcgttaacatctcaaaaagaagct 798
QY 829 gttgaattaacagcagtaacagtgaaacggaacacataaatacaacttcagcgtgcgaa 888
Db 799 gttgaattaacagcagtaacagtgaaacggaacacataaatacaacttcagcgtgcgaa 858
QY 889 aatgaagaatacagtaaaaactgaataactgtacgtatctataaagtgcgaaacgttaacatcca 948
Db 859 aatgaagaatacagtaaaaactgaataactgtacgtatctataaagtgcgaaacgttaacatcca 918
QY 949 ttggcccttaatacgggaagtgtttcttactacacgaacggaaaactatcacctgtgagat 1008
Db 919 ttggcccttaatacgggaagtgtttcttactacacgaacggaaaactatcacctgtgagat 978
QY 1009 gttccaactccatcgcgaataataatacaggaatacaaaagtaagtaagtaaaagctatcaaaagac 1068
Db 979 gttccaactccatcgcgaataataatacaggaatacaaaagtaagtaagtaaaagctatcaaaagac 1038
QY 1069 aaaaatgcgaagaatcttaagaagaatgcgaatccacttcaagcttcgaaatgtaatgctgta 1128
Db 1039 aaaaatgcgaagaatcttaagaagaatgcgaatccacttcaagcttcgaaatgtaatgctgta 1098
QY 1129 gttaccacaaagtgtttggaaactaatatgtaacaaacacactctctgtaaaacttaagcagcaggt 1188
Db 1099 gttaccacaaagtgtttggaaactaatatgtaacaaacacactctctgtaaaacttaagcagcaggt 1158
QY 1189 acttccgacacgtgacggaatcttacaagtagtaattgtaagaattgtttagacccgtgaacat 1248
Db 1159 acttccgacacgtgacggaatcttacaagtagtaattgtaagaattgtttagacccgtgaacat 1218
QY 1249 gtaaacacgtcgaacgttactatcatcagatgttgaaactggaacacgtatccagatatt 1308
Db 1219 gtaaacacgtcgaacgttactatcatcagatgttgaaactggaacacgtatccagatatt 1278
QY 1309 gtaacttaactctgtttctaataactatacttaacgtttaaagaagcgttagtaactgtgtaaa 1368
Db 1279 gtaacttaactctgtttctaataactatacttaacgtttaaagaagcgttagtaactgtgtaaa 1338
QY 1369 caataataaacttgtaatacaataatgttaaaacatttaactgtttacatgcagaaagcttacc 1428
Db 1339 caataataaacttgtaatacaataatgttaaaacatttaactgtttacatgcagaaagcttacc 1398
QY 1429 gaattagtgcttaccatgcgaacgcgaatcagacacaaactgttgtaacgcttccatacttta 1488
Db 1399 gaattagtgcttaccatgcgaacgcgaatcagacacaaactgttgtaacgcttccatacttta 1458

QY 1489 ggtgtacaaacttatatctactgggttctcttacaacaaatgtttcgggttaaatgtgctggt 1548
Db 1459 ggtgtacaaacttatatctactgggttctcttacaacaaatgtttcgggttaaatgtgctggt 1518
QY 1549 ggtgtacaaacttatatctactgggttctcttacaacaaatgtttcgggttaaatgtgctggt 1608
Db 1519 ggtgtacaaacttatatctactgggttctcttacaacaaatgtttcgggttaaatgtgctggt 1578
QY 1609 aagttagaacgaattacttactgactgataactttgtatagttgaaaaagaatcgtgtaca 1668
Db 1579 aagttagaacgaattacttactgactgataactttgtatagttgaaaaagaatcgtgtaca 1638
QY 1669 gttgtgtctctgaacaaataataatgcgaacgttaaaatgttaaaatgtttgaatgtccaaa 1728
Db 1639 gttgtgtctctgaacaaataataatgcgaacgttaaaatgtttgaatgtttgaatgtccaaa 1698
QY 1729 ggcgaacttaagaanaataacatcatcaaatcaaaattaaaaagcttgaaagtcagat 1788
Db 1699 ggcgaacttaagaanaataacatcatcaaatcaaaattaaaaagcttgaaagtcagat 1758
QY 1789 aaggttatgtaataggaactgttaacgagaanaacatagatctcaaaactcgaactta 1848
Db 1759 aaggttatgtaataggaactgttaacgagaanaacatagatctcaaaactcgaactta 1818
QY 1849 actgtctctacagtattatagcgttaagctctaaaaatgtagcagcgtgattaaaaagtaact 1908
Db 1819 actgtctctacagtattatagcgttaagctctaaaaatgtagcagcgtgattaaaaagtaact 1878
QY 1909 gaagctcaaaatattactgtgaaagttctcagagaatttaaatcaatttaatgtatcacac 1968
Db 1879 gaagctcaaaatattactgtgaaagttctcagagaatttaaatcaatttaatgtatcacac 1938
QY 1969 gttcgggtgaac 2028
Db 1939 gttcgggtgaac 1998
QY 2029 tctgctcttacaac 2088
Db 1999 tctgctcttacaac 2058
QY 2089 ggaacatacaaaagtgaagaagtctgctgtaaaccaattgaagaagcttaacacaaagcttca 2148
Db 2059 ggaacatacaaaagtgaagaagtctgctgtaaaccaattgaagaagcttcaacaaagcttca 2118
QY 2149 gtagtggttcggttaaaagctgtaacacacacacacacacacacacacacacacacacacac 2208
Db 2119 gtagtggttcggttaaaagctgtaacacacacacacacacacacacacacacacacacacac 2178
QY 2209 gcaactaacatattctataactatcaactgaaagctgaagaagcttaacacaaagcgttca 2268
Db 2179 gcaactaacatattctataactatcaactgaaagctgaagaagcttcaacacaaagcgttca 2238
QY 2269 acaaaagatctcaaaagctgtaattcttaaaagaacgcgtgaatgcgttacttaacttaacgaac 2328
Db 2239 acaaaagatctcaaaagctgtaattcttaaaagaacgcgtgaatgcgttacttaacttaacgaac 2298
QY 2329 gttgtagcaggttcgaanaattcaactatccaaatttagcgaagaatttaaaacttctagtggt 2388
Db 2299 gttgtagcaggttcgaanaattcaactatccaaatttagcgaagaatttaaaacttctagtggt 2358
QY 2389 tcttttagtggttgcaaaagtaactgttcgaagaatttaaaacttctagtggt 2448
Db 2359 tcttttagtggttgcaaaagtaactgttcgaagaatttaaaacttctagtggt 2418
QY 2449 gttacttggaacaaacttatcagttgctccaaagaacgttcaaatgttaagtaaacgcgt 2508
Db 2419 gttacttggaacaaacttatcagttgctccaaagaacgttcaaatgttaagtaaacgcgt 2478
QY 2509 gctgtgttacaataactgcttctgaacataacgaacaaagctgcgaatgtgccttggtta 2568
Db 2479 gctgtgttacaataactgcttctgaacataacgaacaaagctgcgaatgtgccttggtta 2538

QY	2563	qtaqgtaaqgctctctactcgaagagatgttcgagctgcgctgaagctgaatgtaataatgtaataa	2563
Dh	2539	gtaagtaaqgctctctactctgtaagatgttcgagctgcgctgaagctgaatgtaataagataaa	2538
QY	2639	gatatlttaaatctctgttcaaaagatctgtagaacaacactgttagctcttcgtgagaagcttgct	2658
Dh	2539	gtaagtaaqgctctctactctgtaagatgttcgagctgcgctgaagctgaatgtaataagataaa	2538
QY	2689	gcaaaagatgtgtcaaaacgctcttcgcgcgaattcccaaaagcaacgcgcgaattgtatcaact	2748
Dh	2659	gacaaagatgtgtcaaaacgctcttcgcgcgaattcccaaaagcaacgcgcgaattgtatcaact	2718
QY	2749	aagagctbatatgtatgtcaattccaaatgtaaactgtattcgcggaagctttaaactgtgaacatc	2808
Dh	2719	aagagctbatatgtatgtcaattccaaatgtaaactgtattcgcggaagctttaaactgtgaacatc	2778
QY	2809	gtctgttaaaagctgtcgaacgagatatacgctgatacgctcgttaactgttaacagcatatagaagctct	2868
Dh	2779	gtctgttaaaagctgtcgaacgagatatacgctgatacgctcgttaactgttaacagcatatagaagctct	2838
QY	2869	acaataataattgtcatccaactccatcccaagaattaaagctctgtatcaagctgttaactgtta	2928
Dh	2839	acaataataattgtcatccaactccatcccaagaattaaagctctgtatcaagctgttaactgtta	2898
QY	2929	acaattgtacggtgtcgtagagataaagtaagttaacacacatctctaaabacaattactctgttc	2988
Dh	2859	acaattgtacggtgtcgtagagataaagtaagttaacacacatctctaaabacaattactctgttc	2958
QY	2989	aagactgtatctgtgaattcccaagcttatcttcaatacaagattgtgcgaagctgtgcgaattac	3048
Dh	2959	aagactgtatctgtgaattcccaagcttatcttcaatacaagattgtgcgaagctgtgcgaattac	3018
QY	3049	gtctgaacggttctcaaaacaattatcaactgtaaattcaagatattccaaagcttccaaacccaacatc	3108
Dh	3019	gtctgaacggttctcaaaacaattatcaactgtaaattcaagatattccaaagcttccaaacccaacatc	3078
QY	3109	aactcctaagaaggtgtgcgcggaactcttaattacaaattacacacttttgtaattgtaataat	3168
Dh	3079	aactcctaagaaggtgtgcgcggaactcttaattacaaattacaaacttttgtaattgtaataat	3138
QY	3169	gaaaataaaacaatacaaaatctgttatccaacaagtgtaacaactgtgaagatgttcta	3228
Dh	3139	gaaaataaaacaatacaaaatctgttatccaacaagtgtaacaactgtgaagatgttcta	3198
QY	3229	tatgaactgaagcagttctccaagaagtcttcaaaacgcgttaactgtatctgaatgcgaagttaca	3288
Dh	3199	tatgaactgaagcagttctccaagaagtcttcaaaacgcgttaactgtatctgaatgcgaagttaca	3258
QY	3289	tctcatcaacaggttctgtgttgaattcgcgaagataaaacgcgcctatgtatggcggttggtcca	3348
Dh	3259	tctcatcaacaggttctgtgttgaattcgcgaagataaaacgcgcctatgtatggcggttggtcca	3318
QY	3349	tgtgaatgtgaacaaagctatactactcgaagatgtcgtcgaagcaacgcgaactcgcgtctgtatgcctac	3408
Dh	3319	tgtgaatgtgaacaaagctatactactcgaagatgtcgtcgaagcaacgcgaactcgcgtctgtatgcctac	3378
QY	3409	tctcgttcgcgagccagtttgcgccttcaattctcaagaagatgtatctgaatttaacgaatgtcaact	3468
Dh	3379	tctcgttcgcgagccagtttgcgccttcaattctcaagaagatgtatctgaatttaacgaatgtcaact	3438
QY	3469	gtcgcacagttaaaacaaatattatctgtatgtataaaacgcgtgtgaagttatttccaanaagagatgtta	3528
Dh	3439	gtcgcacagttaaaacaaatattatctgtatgtataaaacgcgtgtgaagttatttccaanaagagatgtta	3498
QY	3529	gagcgacagatccaatgtatgtcgaaggtgtctactaaagagacattatgaattaaacacagttactct	3588
Dh	3499	gagcgacagatccaatgtatgtcgaaggtgtctactaaagagacattatgaattaaacacagttactct	3558
QY	3589	ttaatgaattctgaataacagccaagactttataagattgtttgaagtggtgaaggttaagatgtgcga	3648
Dh	3559	ttaatgaattctgaataacagccaagactttataagattgtttgaagtggtgaaggttaagatgtgcga	3618
QY	3649	ggtataatgttcgagataactatatacttactatataagtaa	3687

[illegible]

Query Match 97.5%; Score 3593.4; DB 20; Length 4988;
Best local similarity 99.8%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 85 gaagcggctacagatgtagcaacagtaglaagcaagcaaaaagcaacagctcaaaaaagca 144
D 1249 gtagctcgctacagatgtagcaacagtaglaagcaagcaaaaagcaacagctcaaaaaagca 1308
QY 145 tactatacttaacgacataacagtaacgaaacgtgtgaatctccaaacataacgtcgtga 204
D 1309 tactatacttaacgacataacagtaacgaaacgtgtgaatctccaaacataacgtga 1368
QY 205 taccgtgaatacacaacaagcgaaaaaaacgaataccgtgtgaatcggtagagattgaataa 264
D 1369 tagcggaatacacaacaagcgaaaaaaacgaataccgtgtgaatcggtagagattgaataa 1428
QY 265 gcaggtggtcgcgaaaaaaagcgcttaacttagctgattacaanaaagaaatagaacattac 324
D 1429 gcaggtggtcgcgaaaaaaagcgcttaacttagctgattacaanaaagaaatagaacattac 1488
QY 325 gttttcaaaagcaaaccttaaatctgtggaaagctccgtgtgaacattacatcgatgcttac 384
D 1489 gttttcaaaagcaaaccttaaatctgtggaaagctccgtgtgaacattacatcgatgcttac 1548
QY 385 aactatgcaacaaatctagaacgaatgccccaaagctagaagctgtgttcacaagcaaaa 444
D 1549 aactatgcaacaaatctagaacgaatgccccaaagctagaagctgtgttcacaagcaaaa 1608
QY 445 gatttgaanaaagcagaaacaatactacacaanaattccttaatgaataaactcgcgaca 504
D 1609 gatttgaanaaagcagaaacaatactacacaanaattccttaatgaataaactcgcgaca 1668
QY 505 gtcaattttgatctgcgtatatagtttaaaaacaactcgtgtgaatttactcgccttaacattaaa 564
D 1669 gtcaattttgatctgcgtatatagtttaaaaacaactcgtgtgaatttactcgccttaacattaaa 1728
QY 565 gcaaaaagcacaagaactcgcgcgaacagcttaattatgaatacttcacgttgcgaatgaagcg 624
D 1729 gcaaaaagcacaagaactcgcgcgaacagcttaattatgaatacttcacgttgcgaatgaagcg 1788
QY 625 cgcgaagctacaagacgctgtgaaagcagcgcaatttagacaanaagctaaagctgcgtgtgat 684
D 1789 cgcgaagctacaagacgctgtgaaagcagcgcaatttagacaanaagctaaagctgcgtgtgat 1848
QY 685 caaatcaatacaatacttaaccaaaaagtaacagatgcttcaaaaactgaaccaacagaagta 744
D 1849 caaatcaatacaatacttaaccaaaaagtaacagatgcttcaaaaactgaaccaacagaagta 1808
QY 745 gcaaaaaaagcatctagaatgtagaagctgcgtcttaactccaaaagttagaagtgtagatg 804
D 1909 gcaaaaaaagcatctagaatgtagaagctgcgtcttaactccaaaagttagaagtgtagatg 1868
QY 805 ggcattacaactcaaaaacaaagcgtgtgaatlaaacagcaatgacagtggaacggaacacta 864
D 1969 ggcattacaactcaaaaacaaagcgtgtgaatlaaacagcaatgacagtggaacggaacacta 2028
QY 865 aaattcaaatcttcagctgtgtgaaatgaagatacagtaaaacgttaaacctgtacgtatc 924
D 2029 aaattcaaatcttcagctgtgtgaaatgaagatacagtaaaacgttaaacctgtacgtatc 2088
QY 925 tataaaggtgaacggttaaacatttcgaccttaatacggcagagattttttttatccaca 984
D 2089 tataaaggtgaacggttaaacatttcgaccttaatacggcagagattttttttatccaca 2148
QY 985 gacgggaaaaacttcaactgtgtgagctgttcaactccattcgaataataatacggagtataaa 1044
D 2149 gacgggaaaaacttcaactgtgtgagctgttcaactccattcgaataataatacggagtataaa 2208
QY 1045 gtaagtgattaaaggtatttaaaagcaaaaatgtgcaaaagaaatttaaaagaaatgtcaattcact 1104
D 2209 gtaagtgattaaaggtatttaaaagcaaaaatgtgcaaaagaaatttaaaagaaatgtcaattcact 2268

QY 1105 ttcaagcttcgaaatgtagctgtgtagtactcaagtglttggaaactaattgaacaaacac 1164
D 2269 ttcaagcttcgaaatgtagctgtgtagtactcaagtglttggaaactaattgaacaaacac 2328
QY 1165 acttcgttaaaacttagcagcaggtactcttcgacactgaacgaatacttaacagtagatttt 1224
D 2329 acttcgttaaaacttagcagcaggtactcttcgacactgaacgaatacttaacagtagatttt 2388
QY 1225 gataaagtgtgtaacacgtgaagaaactgtgaacagctgaaagttactttcaagtglttga 1284
D 2389 gataaagtgtgtaacacgtgtgaagaaactgtgaacagctgaaagttactttcaagtglttga 2448
QY 1285 actggaanaagcattcccaatgtgcaactctcgtgtgtctcaaatctactatacgtta 1344
D 2449 actggaanaagcattcccaatgtgcaactctcgtgtgtctcaaatctactatacgtta 2508
QY 1345 aaagaagcgttgaatgaactgtgtaaacacataaactgtgtcaatcaaatgtttaaacatga 1404
D 2509 aaagaagcgttgaatgaactgtgtaaacacataaactgtgtcaatcaaatgtttaaacatga 2568
QY 1405 actggtttacaatgcaagaagcttaacggttagtggtttacaacgaacgaatcaagaacact 1464
D 2569 actggtttacaatgcaagaagctttacggttagtggtttacaacgaacgaatcaagaacact 2628
QY 1465 gttgctacccgtctcctactactattaggtggtgtacaaactttactcgtgtctctacaaca 1524
D 2629 gttgctacccgtctcctactactattaggtggtgtacaaactttactcgtgtctctacaaca 2688
QY 1525 aatgtttggtgttaaatgttgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1584
D 2689 aatgtttggtgttaaatgttgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2748
QY 1585 caattcaacaacaacglttgtaactaaggttagagaagaatcctaacttagctgaataccttgta 1644
D 2749 caattcaacaacaacglttgtaactaaggttagagaagaatcctaacttagctgaataccttgta 2808
QY 1645 ttactgtgaanaaagaactcgtgtacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1704
D 2809 ttactgtgaanaaagaactcgtgtacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2868
QY 1705 aaaaatgttaacttttagtgccaaaagcggaaccttaaaagaaatacaatctatacaatcaaa 1764
D 2869 aaaaatgttaacttttagtgccaaaagcggaaccttaaaagaaatacaatctatacaatcaaa 2928
QY 1765 attaaaaaaggtcttgaaagctccgaataaggtatagaattagcactgtttaacgagaanaaa 2988
D 2929 attaaaaaaggtcttgaaagctccgaataaggtatagaattagcactgtttaacgagaanaaa 3048
QY 1825 tatgaattcaaaaactgaagacttaactgcctccaaagcttatagcggttaacggtcttaaaat 1884
D 2989 tatgaattcaaaaactgaagacttaactgcctccaaagcttatagcggttaacggtcttaaaat 3048
QY 1885 ggcgaacgctgtatataaanaatgaacgtgaacgtcaagaaattactgtgaaggtctccaaagaaat 1944
D 3049 ggcgaacgctgtatataaanaatgaacgtgaacgtcaagaaattactgtgaaggtctccaaagaaat 3108
QY 1945 ttaaaatacaattcaatgctacaacgcttttcggttagcagaacatacaacaggtccaagttgct 2004
D 3109 ttaaaatacaattcaatgctacaacgcttttcggttagcagaacatacaacaggtccaagttgct 3168
QY 2005 gtagttaaagcggtgtgcaaacctatctgctgttcaagcaagtggaacgtcttaacagttagt 2064
D 3169 gtagttaaagcggtgtgcaaacctatctgctgttcaagcaagtggaacgtcttaacagttagt 3228
QY 2065 gttgaagcggttactgtgtcaaatgtgaacatacaaaagtgaagttgtctgtcttaaacatata 2124
D 3229 gttgaagcggttactgtgtcaaatgtgaacatacaaaagtgaagttgtctgtcttaaacatata 3288
QY 2125 gaacgttaaaccaaggttaacaataattagtaggtgtcgtgtcaaaagtgcaaacgctccgtgttaa 2184
D 3289 gaacgttaaaccaaggttaacaataattagtaggtgtcgtgtcaaaagtgcaaacgctccgtgttaa 3348
QY 2185 gatgctgtcaaatgcaaatatttagcaactaactatatactatacatttaacactgtgaagt 2244

```

Db 3349 gatgctgcaaatgcaaatcttagcaactaacatatactatacttaacattcaaacggaaggt 3408
Qy 2245 caagaagtaaacgacacacacaggtttacacaaagtatacaaggtgttttttaaaagagct 2304
Db 3409 caagaagtaaacgacacacacaggtttacacaaagtatacaaggtgttttttaaaagagct 3468
Qy 2305 gatgcaagtaactacacttaacgacgttgaacgaggttaacaaatcacatccaaattacg 2364
Db 3469 gatgcaagtaactacacttaacgacgttgaacgaggttaacaaatcacatccaaattacg 3528
Qy 2365 gaagaatataaaactcttagtcttcttagtgggtgagcaagtaactgttcggaataat 2424
Db 3529 gaagaatataaaactcttagtcttcttagtgggtgagcaagtaactgttcggaataat 3588
Qy 2425 acaaaacagagtaggtgaggtgtcgtgacacggaacacatgatacgtgtccttaagaca 2484
Db 3589 acaaaacagagtaggtgaggtgtcgtgacacggaacacatgatacgtgtccttaagaca 3648
Qy 2485 gatgcaaatgttaaaagttaaacagctgctgtgttaacatgaactgttctgaacataacgac 2544
Db 3649 gatgcaaatgttaaaagttaaacagctgctgtgttaacatgaactgttctgaacataacgac 3708
Qy 2545 aaagatgcaaatgtgctcgtgtgtagtaagaatctctactatgtagtaaatgtcgtatgta 2604
Db 3709 aaagatgcaaatgtgctcgtgtgtagtaagaatctctactatgtagtaaatgtcgtatgta 3768
Qy 2605 gctggtatgttaatttaaggaanaagatatttaattctgtaacaaacgcttctgtacatcca 2664
Db 3769 gctggtatgttaatttaaggaanaagatatttaattctgtaacaaacgcttctgtacatcca 3828
Qy 2665 gtagctctgtgaaagctgtcgtgacaaagatgtgtcaaaacgcttctgtacatcca 2724
Db 3829 gtagctctgtgaaagctgtcgtgacaaagatgtgtcaaaacgcttctgtacatcca 3888
Qy 2725 acaagcctgtcaattgtttacacaaagcttattatgttgatccaatgaactgattta 2784
Db 3889 acaagcctgtcaattgtttacacaaagcttattatgttgatccaatgaactgattta 3948
Qy 2785 ggcgaagtttaaacctgtgaaacatcgtttgttaaaagtacagagtaatgctgtcgtgt 2844
Db 3949 ggcgaagtttaaacctgtgaaacatcgtttgttaaaagtacagagtaatgctgtcgtgt 4008
Qy 2845 actgttaacagcattagaacggtttctacaataaattgttatctactcattcacaagaatta 2904
Db 4009 actgttaacagcattagaacggtttctacaataaattgttatctactcattcacaagaatta 4068
Qy 2905 aaagctgttcaagttctctcgttaacaattgacggtgtgagagataagtagtaacaca 2964
Db 4069 aaagctgttcaagttctctcgttaacaattgacggtgtgagagataagtagtaacaca 4128
Qy 2965 atctctaatacattactcttctcaagactgtatctgcaatcccaacgtttatcttcaatc 3024
Db 4129 atctctaatacattactcttctcaagactgtatctgcaatcccaacgtttatcttcaatc 4188
Qy 3025 agcatgtgacaggtgtgacgttaacggttgcgttcaaaaacaaatgaattgaattcagc 3084
Db 4189 agcatgtgacaggtgtgacgttgaacggttgcgttcaaaaacaaatgaattgaattcagc 4248
Qy 3085 gattcagttccaaacacaaacaaatcacctcttaagaagcgtgcggaacttcaattcaat 3144
Db 4249 gattcagttccaaacacaaacaaatcacctcttaagaagcgtgcggaacttcaattcaat 4308
Qy 3145 taacacttataataatgtataaatgtataaataaacaataaataatgttatcccaaaaggt 3204
Db 4309 taacacttataataatgtataaatgtataaataaacaataaataatgttatcccaaaaggt 4368
Qy 3205 gtaaacacttgacaggttactactcaataaggttagcagtttcaaaaagtttcaaacgtgt 3264
Db 4369 gtaaacacttgacaggttactactcaataaggttagcagtttcaaaaagtttcaaacgtgt 4428
Qy 3265 actgatatgtatagcaaaagttacatcaacaggttcctgttactgacgaaagtaaa 3324

```

```

Db 4429 actgatatgtatagcaaaagttacatcaatcacaaggttcctgttactgacgaaagtaaa 4488
Qy 3325 cctgcttaagtaggccttggttgattgaaatggaacaaagctatactcaggtgtgagca 3384
Db 4489 cctgcttaagtaggccttggttgattgaaatggaacaaagctatactcaggtgtgagca 4548
Qy 3385 acacagactcgtctctttagctgaactcctgttgcggagccaggttccttcaattcaga 3444
Db 4549 acacagactcgtctctttagctgaactcctgttgcggagccaggttccttcaattcaga 4608
Qy 3445 ggtatcagtttaacgaaatgcaactgtgacagtaacaaatattctatgataaactgt 3504
Db 4609 ggtatcagtttaacgaaatgcaactgtgacagtaacaaatattctatgataaactgt 4668
Qy 3505 gaagttattcaaaaagagtagtagacgcagacacatgattgacgtgttactaaagagaca 3564
Db 4669 gaagttattcaaaaagagtagtagacgcagacacatgattgacgtgttactaaagagaca 4728
Qy 3565 ttgtaattacaacagttactcctttagtaactgtgataacagcaagacttaagattgt 3624
Db 4729 ttgtaattacaacagttactcctttagtaactgtgataacagcaagacttaagattgt 4788
Qy 3625 gtaagtagagtttaaaagttagcagagtaattgttgcagatactatactataatgaag 3684
Db 4789 gtaagtagagtttaaaagttagcagagtaattgttgcagatactatactataatgaag 4848
Qy 3685 taa 3687
Db 4849 taa 4851

RESULT 6
ID T05868 standard; DNA; 3399 BP.
AC T05868;
DT 14-APR-1996 (first entry)
XX
DE Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.
XX
KW Chicken leucocytozoan; immunogen; recombinant vaccine; protection;
XX Immunisation; vaccination; ss.
XX Chicken leucocytozoan.
XX
FH Key Location/Qualifiers
FT CDS 1..3399
FT /tag= a
FT misc_feature 1150..3218
FT /tag= b
FT FT "Fragment referred to in the claims, for use as insert in a recombinant vaccine against chicken leucocytozoan disease"
FT FT
XX
XX JP07284392-A.
XX
XX 31-OCT-1995.
XX
XX 19-APR-1994; 94UP-0080643.
XX
XX 19-APR-1994; 94UP-0080643.
XX
XX
XX (DOBU-) DOBUTSUO SEIBUTSUGAKURPEKI SEIZAI KYOKAI.
XX (KITA) KITASATO KENKUSHO SH.
XX
XX WPI; 1996-006311/01.
XX P-PSDB; R97866.
XX
XX Chicken leucocytozoan immunogenic protein - used in a recombinant
XX vaccine against chicken leucocytozoan disease
XX
PS Claim 6; Page 6-9; 35pp; Japanese.

```

XX 705868 encodes a chicken leucocytozoan immunogenic protein, this DNA
CC or a fragment of it can be used in a recombinant vaccine to immunise
CC against chicken leucocytozoan disease. The DNA is used in a vector
CC and operatively linked to an expression regulatory sequence as in
CC standard practice.

50 Sequence 3359 BP; 1577 A; 508 C; 798 G; 516 T; 0 other;

Query Match	2.18;	Score 77.4;	DB 17;	Length 3399;
Best Local Similarity	43.08;	Pred. No. 8.9e-08;		
Matches 436;	Conservative 0;	Mismatches 576;	Indels 3;	Gaps 1

[illegible][illegible]

RESULT

ID Q87587 standard; DNA; 1686 BP.

AC Q87587;

DT 19-DEC-1995 (first entry)

DE DNA encoding Leucocytozoan protozoa structural protein epitope.

KM leucocytozoan protozoa; structural protein; epitope; vaccine; fowl
KM leucocytozanosis; treatment; ss.

05 Leucocytozoan protozoa sp.

PN JP07089995-A.

PD 04-APR-1995.

PE 10-SEP-1993; 93JP-0226078.

PR 10-SEP-1993; 93JP-0226078

PA (DOB-) DOBUTSUYO SEIBUTSUGAKUTEKI SEIZAI KYOKAI
PA (DOB-) DOBUTSUYO SEIBUTSUGAKUTEKI SEIZAI KYOKAI

XX
DB WPT. 1005-167353/33

DR P-PSDB; R70491.

Immune inducing
in production of

PT	LOWL.
YY	

Claim 1; page 12-14; 20pp; Japanese.

CC Q87587-89 encode polypeptides having a whole or partial epitope of a
CC structural protein of *Leishmania* protozoa (see P70491-93). The

polypeptides and DNA encoding them are useful in the prevention of the treatment of leucocytosis of fowl

Query Match	1.73;	Score 62.4;	DB 16;	Length 1686;
Best Local Similarity	45.33;	Pred. No. 0.00013;		
Matches 308;	Conservative 0;	Mismatches 366;	Indels 6;	Gaps 2

QY	96	agatgtagcacacgtatgataagcccaagcaaaagcacagttcaaaaaagcatataactta	155
Db	141	aggaagaaacaaagacacagagagagacacagacagaaatctgtgaagaaacaaacacagatga	200
QY	156	cagcgatacacgtatgaacgtgtgtgtatctccaaacataacacgtatgatgtgtata	215
Db	201	agagagacacaaagaaagagagagatgtgaagaaagaaagaaagaaagaaagaaagaaagaa	260
QY	216	caacaaacgagaaaaaaacgataccgtgaatgcgtgtacatagtgagataaagcaggttgcgc	275
Db	261	agaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa	320
QY	276	gaaaaaaaacgcttaactagctgtgtttcaaaaaagatatgaacctaacgtttccaagc	335
Db	321	aggaagaaacaaacagcttgaaagatgaaagaaagaaagaaagaaagaaagaaagaaagaa	380

CC SCARF gene containing 73 copies of the trinucleotide (CAG) repeat unit
CC (the present sequence) to a substrate, and hybridizing with amplified
CC test genomic DNA containing copies of the trinucleotide units, PCR
CC amplified with labeled primers (X89889-90). The new method is useful for
CC diagnosis of SCARF syndrome, and for determining the severity of the
CC disease. The present sequence represents the SCARF gene fragment
CC containing 73 trinucleotide (TNR) repeats.

SO Sequence 397 BP; 124 A; 104 C; 104 G; 65 T; 0 other;

Query Match 1.5%; Score 56.2; DB 20; Length 397;
Best Local Similarity 47.3%; Pred. No. 0.0019;
Matches 169; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 1454 caccacacaaatgttgcacgcctctactctttagtggtgaacattatctctgtt 1513
DB 358 cccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 299
QY 1514 ctctacacacaaatgttgggtgaatgtgctggtggtgaatgaactggaattatt 1573
DB 298 ctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 239
QY 1574 atcctggtcttcacattcaacaaacgcttctactaagtaagacgaactacttagctg 1633
DB 238 ctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 179
QY 1634 ataacttgatatttagtgtaaaagaatctgtgataagttgtgtctcgaactaaataata 1693
DB 178 ctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 119
QY 1694 atcgagcgctaaatgttaactttagtgccaaagcgacacttaagaataatactct 1753
DB 118 tcggaacattcaaaagtgagatatttttaaaacaaacttaaaagattatataacc 59
QY 1754 atcaaatcaaatatataaagcgctgaagtcgaataaggtattgaattgagcactg 1810
DB 58 atgagaaacaaactatttcattcaaggaataatcacattgtttcacgaatcaagtagtactg 2

RESULT 12

T42063/C
ID T42063 standard; DNA; 1830121 BP.

XX T42063;

AC T42063;

DT 14-Sep-1999 (first entry)

DE Haemophilus influenzae complete genome sequence.

XX Genome; bacterium; Haemophilus influenzae; computer readable medium;
KM expression modulating fragment; regulation; gene expression; vector;
KW organism; open reading frame; ORF; ds.

OS Haemophilus influenzae.

XX MO9633276-A1.

PN 24-OCT-1996.

PD 22-APR-1996; 96MO-U0505320.

PR 07-JUN-1995; 95US-0487429.

PR 21-APR-1995; 95US-0426787.

PR 07-JUN-1995; 95US-0476102.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (UJUD) UNIV JOHNS HOPKINS.

XX Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;
XX WPI; 1996-485782/48.

PT Haemophilus influenzae Rd genome recorded on computer readable
PT medium - useful for identifying commercially important nucleic acid
PT fragments by homology searching

PS Claim 1; Page 77.2-77.1091; 1291pp; English.

XX This sequence represents the complete genome sequence of the bacterium
CC Haemophilus influenzae strain Rd. The invention relates to a computer
CC readable medium (CRM) having recorded upon it the complete H influenzae
CC nucleotide sequence (I), a representative fragment of (I) or a nucleotide
CC sequence at least 99% identical to (I). By providing the full-length
CC genomic sequence in a computer readable form, it is possible to identify
CC commercially important nucleic acid fragments and expression modulating
CC fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
CC regulate the expression of a nucleic acid molecule. Vectors and altered
CC organisms comprising the predicted ORFs can be used to produce any of the
CC polypeptide fragments of the H. influenzae Rd genome.

SO Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;

Query Match 1.4%; Score 53; DB 17; Length 1830121;
Best Local Similarity 45.5%; Pred. No. 0.12;
Matches 277; Conservative 0; Mismatches 320; Indels 12; Gaps 2;

QY 601 gatataccgctgaatgaagcgcggaagatgaagacgctgtgaagaggaattta 650
DB 1806740 ggtacgacacggtgacacacacacacacacacacacacacacacacacac 1806691
QY 661 gacaaagctaaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720
DB 1806680 gctattactaaagtgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1806624
QY 721 ttcaaaactgaactaagaagaatgaagaaagccttgatgacatgagctgctt 780
DB 1806623 tttaaatatctgacacacacacacacacacacacacacacacacacacac 1806564
QY 781 atcccaaaagttgaagttgaagttgaagttgaagttgaagttgaagttgaagttga 840
DB 1806563 aatctttaaagaaatgaagttgaagttgaagttgaagttgaagttgaagttga 1806513
QY 841 gcaatcagatgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 900
DB 1806512 gggggcagctgtagggcagctgtagggcagctgtagggcagctgtagggcag 1806453
QY 901 gtaaacgttaatactgtacgtatctataaagtgaagcgttaacattgccttaatt 960
DB 1806452 aaagaaagaaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaag 1806393
QY 961 acggcagatgttcttatctacagcgaagcgaagcgaagcgaagcgaagcgaagcga 1020
DB 1806392 gctacgcttacttccaaatcttaaaagcgttaccatattaccgtttagtgttga 1806333
QY 1021 ttccaaatatacagagatgaagatgaagatgaagatgaagatgaagatgaagatga 1080
DB 1806332 acttaagcgaatagcggctttgaaaaagatgacgatactatgaatgaatgaat 1806273
QY 1081 gaatttaagaagatgcatcacttcaagctcgaatgaatgaatgaatgaatgaatga 1140
DB 1806272 caaaacactgatatgttttaactgttgcgttaattgaatgaatgaatgaatga 1806213
QY 1141 ttggacactaagttaacaaacacacacacacacacacacacacacacacacacac 1200
DB 1806212 ttggaaactgttttaaaacttgagacgactgatacagatccggttaaaactga 1806153
QY 1201 gacgatact 1209
DB 1806152 gctactgct 1806144

RESULT 13

XX X61812 standard; DNA; 1039 BP.
ID X61812

```
XX AC X61812;
XX DE 19-JUL-1999 (first entry)
XX DE B. burgdorferi antigenic protein coding sequence, t5-14.nt.
XX KM Antigenic protein; vaccine; Lyme disease; infection; detection: ss.
XX OS Borrelia burgdorferi.
XX PN W09859071-A1.
XX PD 30-DEC-1998.
XX PF 18-JUN-1998; 98WO-US12718.
XX PR 03-SEP-1997; 97US-0057483.
XX PR 20-JUN-1997; 97US-0050359.
XX PR 22-JUL-1997; 97US-0053344.
XX PR 22-JUL-1997; 97US-0053377.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (MEDI-) MEDIMUNE INC.
XX PI Chol GH, Erwin AL, Hanson MS, Lathigra R;
XX DR WPI, 1999-189980/16.
XX DR P-PSDB; Y20115.
XX PT New isolated Borrelia burgdorferi nucleic acids - used to develop
XX PT products for the diagnosis, prevention and treatment of diseases
XX PT caused by Borrelia, particularly Lyme disease
XX PS Claim 1; Page 202; 275pp; English.
XX CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
XX CC invention, which is suitable for use in a vaccine. The Bb polypeptides
XX CC can be used in vaccines for eliciting protective antibodies to members of
XX CC the Borrelia genus, particularly for the use against Lyme disease in
XX CC humans and animals. They can be used for preventing or attenuating an
XX CC infection caused by a member of the Borrelia genus. The products can also
XX CC be used for detection of members of the Borrelia genus.
XX SQ Sequence 1039 BP; 526 A; 94 C; 212 G; 207 T; 0 other;

Query Match 1.4%; Score 51.8; DB 20; Length 1039;
Best Local Similarity 47.6%; Pred. No. 0.023;
Matches 152; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 177 tgggtgaattcccaacattagatgtatattatgtacacacaaagcgaaaaagata 236
    |||||
DB 280 tgggtgaattcccaacaaagaaagaaagaaatcttaagataagttgaggaaagaaaaa 339
QY 237 ccgtgatgcgtgagcattagatgaataagcaggtgycgcgaaaaaagacgttacttagc 296
    |||||
DB 340 agttgtgagggagaaagaaagaaagacagatagtaaaaaagaaagaaatgtgagaagcaag 399
QY 297 tgatttcaaaaagaatatatgaacttcgttttcaaaagcaaaccttaattctggcgaagc 356
    |||||
DB 400 tcaaaagcaaaaagaaagaaagaaagaaactctaaagaaagaaacaaagaaagaaagc 459
QY 357 tcgtgtgacaaacttaacttcgttcacactatgcaacacaaatagaaagaaatgagcga 416
    |||||
DB 460 aaaaagctagagcagatagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 519
QY 417 agagctagaagcgtcttcaagcaaaagaaatagaaagaaagaaagaaagaaatctatcaca 476
    |||||
DB 520 acaggaagaaagctaggtttaaagcagaaagaaagaaagaaagaaagaaagaaagaa 579
QY 477 aattcttatgaataaag 495
    |||||
```

```
DB 580 acaagaaagaaagaaagaa 598

RESULT 14
ID X61811
ID X61811 standard; DNA; 1125 BP.
XX AC X61811;
XX DE 19-JUL-1999 (first entry)
XX DE B. burgdorferi antigenic protein coding sequence, f5-14.nt.
XX KM Antigenic protein; vaccine; Lyme disease; infection; detection: ss.
XX OS Borrelia burgdorferi.
XX PN W09859071-A1.
XX PD 30-DEC-1998.
XX PF 18-JUN-1998; 98WO-US12718.
XX PR 03-SEP-1997; 97US-0057483.
XX PR 20-JUN-1997; 97US-0050359.
XX PR 22-JUL-1997; 97US-0053344.
XX PR 22-JUL-1997; 97US-0053377.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (MEDI-) MEDIMUNE INC.
XX PI Chol GH, Erwin AL, Hanson MS, Lathigra R;
XX DR WPI, 1999-189980/16.
XX DR P-PSDB; Y20114.
XX PT New isolated Borrelia burgdorferi nucleic acids - used to develop
XX PT products for the diagnosis, prevention and treatment of diseases
XX PT caused by Borrelia, particularly Lyme disease
XX PS Claim 1; Page 202; 275pp; English.
XX CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
XX CC invention, which is suitable for use in a vaccine. The Bb polypeptides
XX CC can be used in vaccines for eliciting protective antibodies to members of
XX CC the Borrelia genus, particularly for the use against Lyme disease in
XX CC humans and animals. They can be used for preventing or attenuating an
XX CC infection caused by a member of the Borrelia genus. The products can also
XX CC be used for detection of members of the Borrelia genus.
XX SQ Sequence 1125 BP; 565 A; 101 C; 223 G; 236 T; 0 other;

Query Match 1.4%; Score 51.8; DB 20; Length 1125;
Best Local Similarity 47.6%; Pred. No. 0.024;
Matches 152; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 177 tgggtgaattcccaacattacatgtatattatgtacacaaagcgaaaaagata 236
    |||||
DB 363 tgggtgaattcccaacaaagaaagaaagaaatcttaagataagttgaggaaagaaaaa 422
QY 237 ccgtgatgcgtgagcattagatgaataagcaggtgycgcgaaaaaagacgttacttagc 296
    |||||
DB 423 agttgtgagggagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaag 482
QY 297 tgatttcaaaaagaatatatgaacttcgtttcacaagcaaaccttaattctggcgaagc 356
    |||||
DB 483 tcaaaagcaaaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaag 542
QY 357 tcgtgtgacaaacttaacttcgttcacactatgcaacacaaatagaaagaaatgagcga 416
    |||||
DB 543 aaaaagctagagcagatagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 602
    |||||
```


